

MEDIEN



INSTANT SEARCH - HSMW_TUC Team

TRECVID 2019

Tony Rolletschke - University of Applied Science Mittweida

12/11/19

Introduction - Team

Our INS approach with pretrained models and web based interactive evaluation by using a management system

Who is hiding behind 'our'?

- Rico Thomanek
- Christian Roschke
- Benny Platte
- Tony Rolletschke
- Tobias Schlosser
- Manuel Heinzig
- Danny Kowerko
- Matthias Vodel
- Frank Zimmer
- Maximilian Eibl
- Marc Ritter



TECHNISCHE UNIVERSITÄT
CHEMNITZ



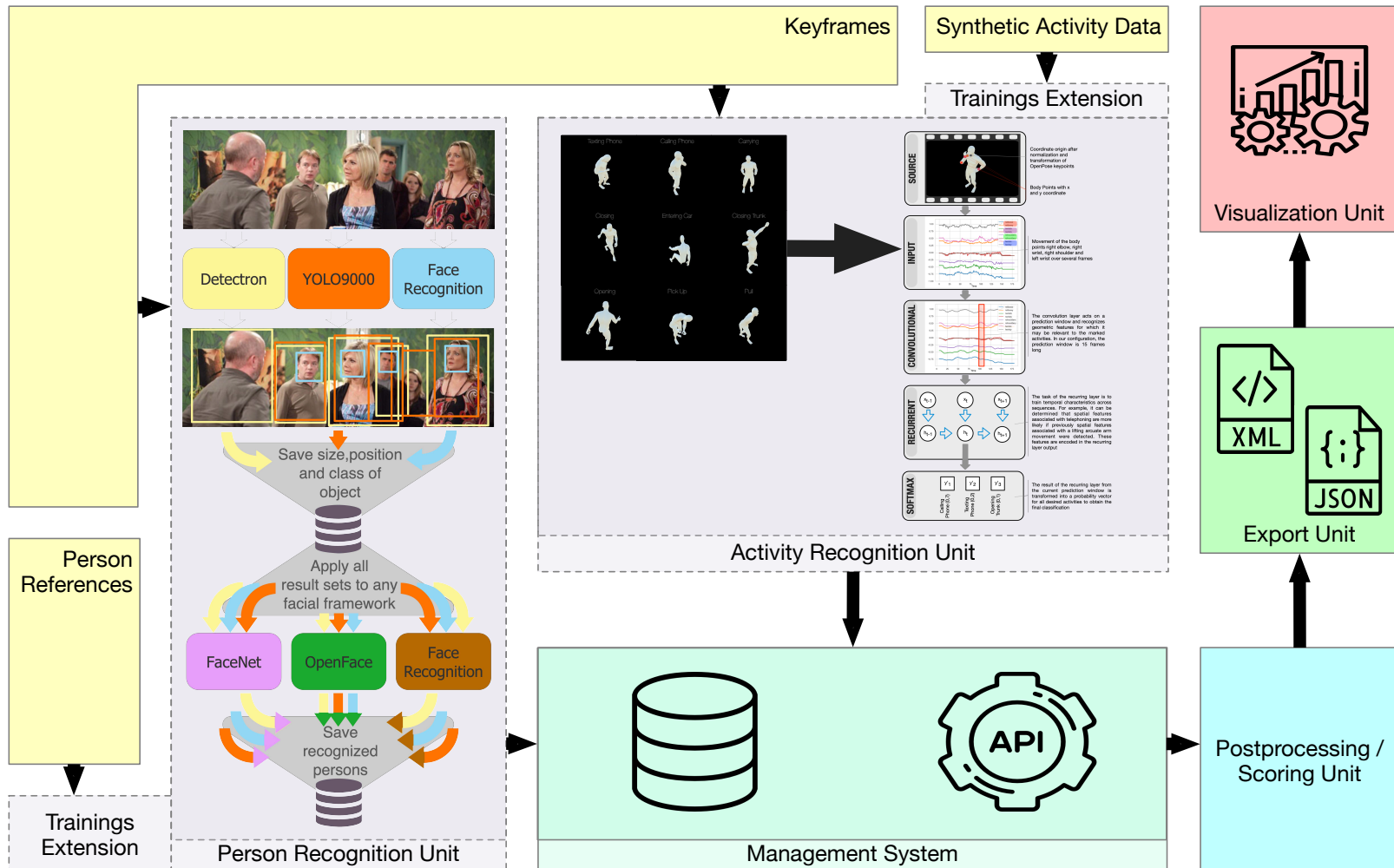
Introduction - Focus

Our INS approach with pretrained models and web based interactive evaluation by using a management system

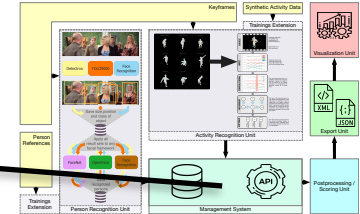
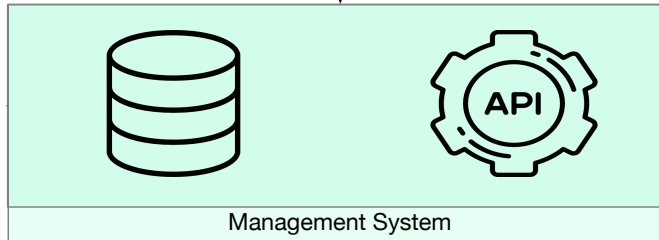
- We mainly wanted to improve and extend our system from last year
- While using the same hardware we could increase the system performance significantly
- That was the result of a new approach with distributed clusters administrated by a management system
- The focus was on high adaptability, which enables processing tasks in different application domains

System - Architecture

Holistic server-client approach



System - Architecture



Database

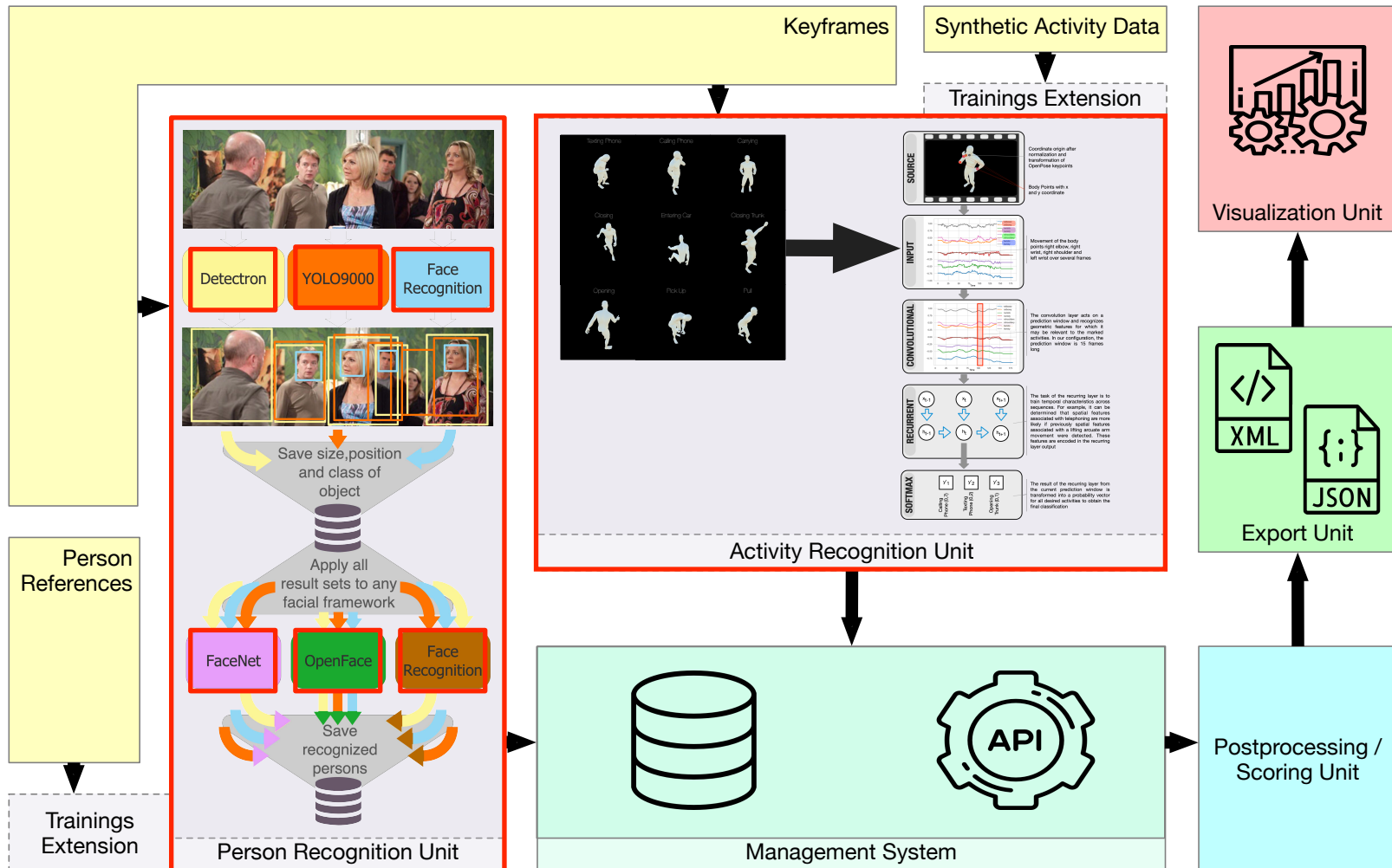
- Handles the persistent storage
- All data can be accessed by an API

Management

- Add or remove processing nodes
- Tasks can be allocated with different nodes
- Automatic task completion
- Error handling

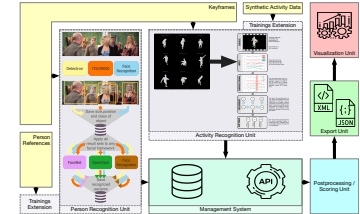
System - Architecture

Distribution with docker



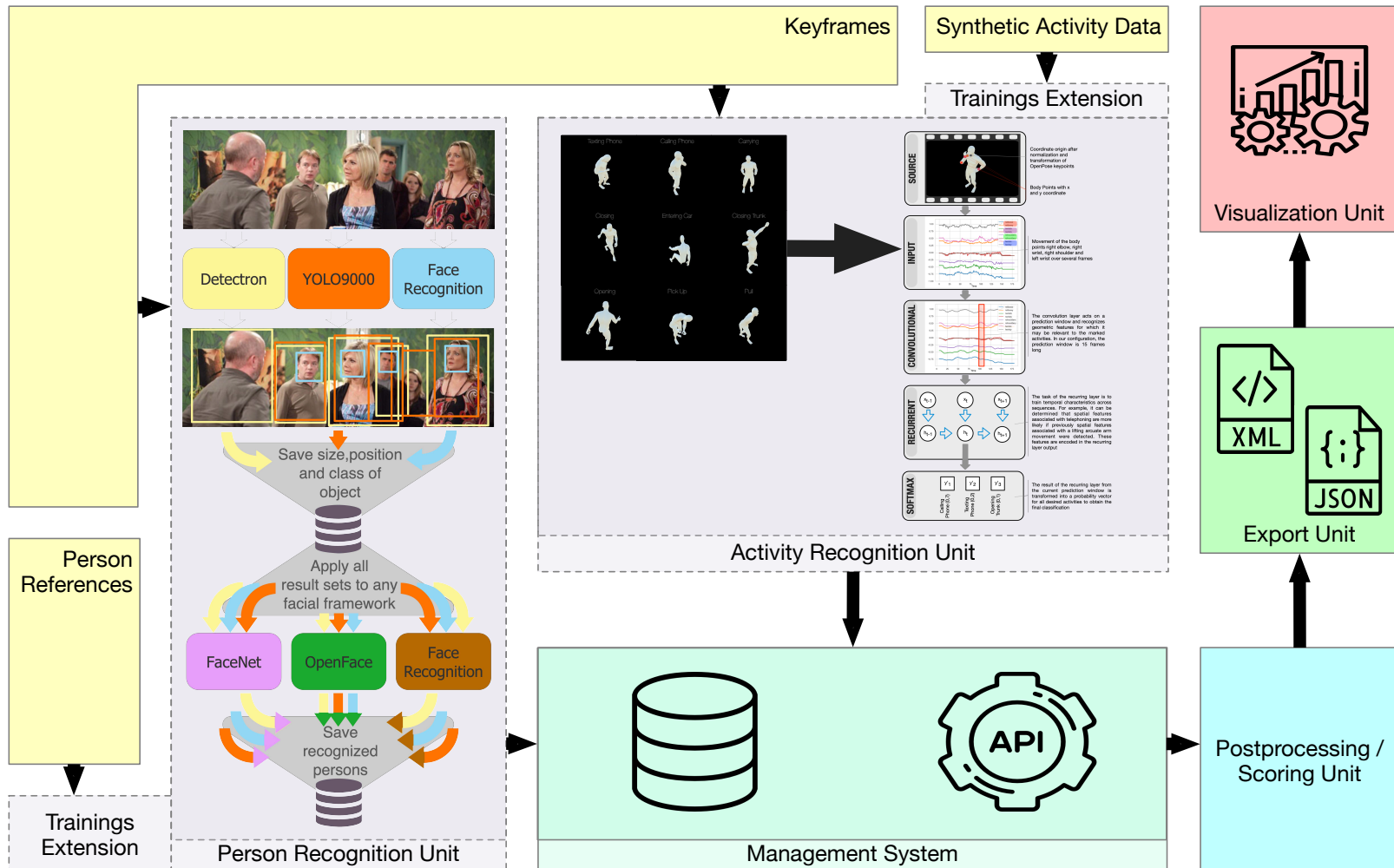
Distribution with docker

- Distribution without manual configuration
- Containers to be run multiple times on each node
- Each host can run on maximum load



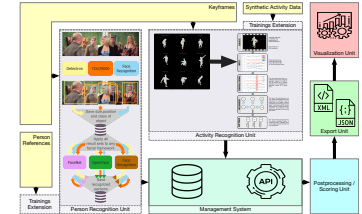
System - Architecture

State-of-the-art frameworks to identify people and objects:



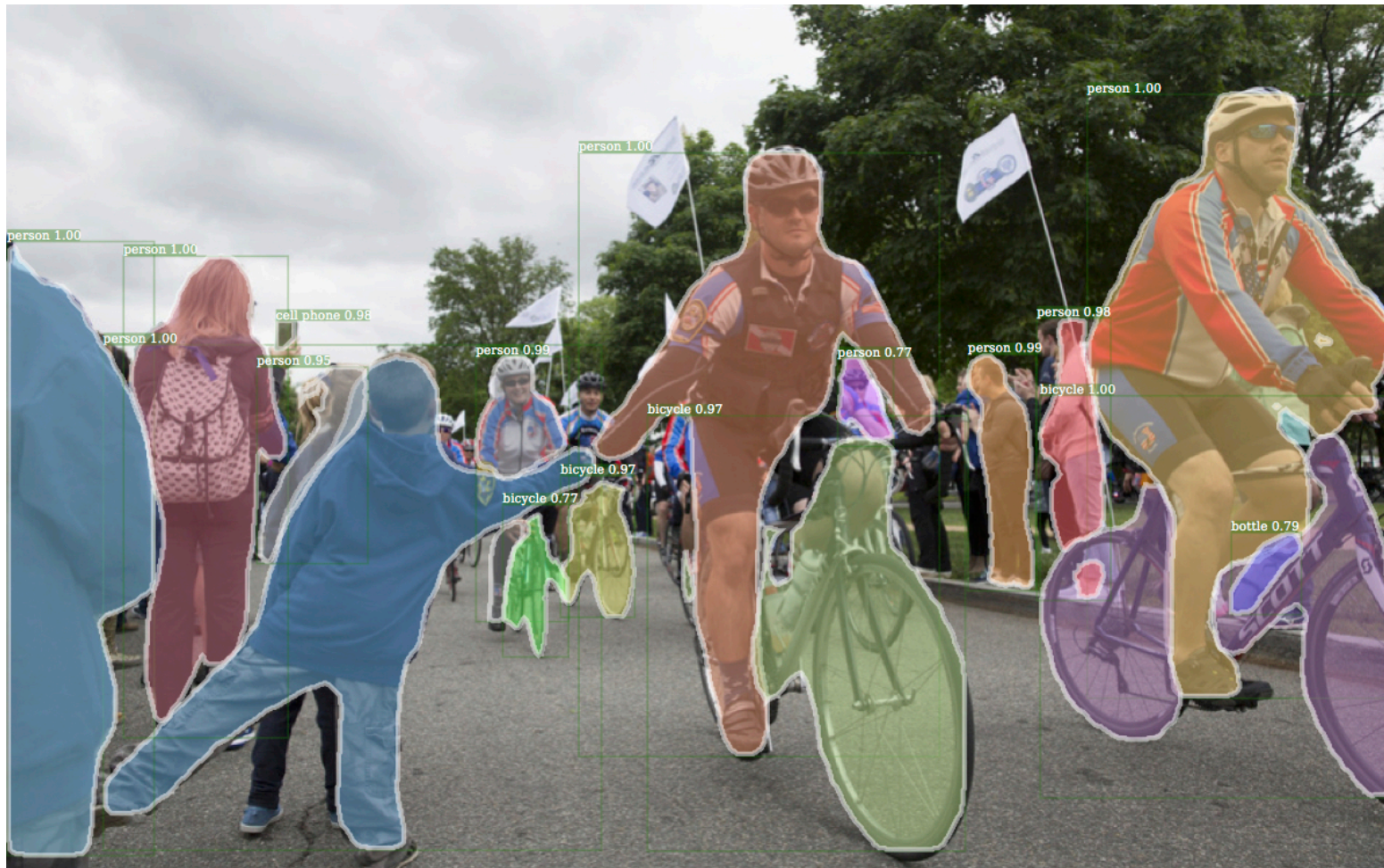
State-of-the-art frameworks to identify people and objects:

- Detectron
- Yolo9000
- FaceNet
- OpenFace
- FaceRecognition
- TuriCreate



System - Frameworks - Detectron

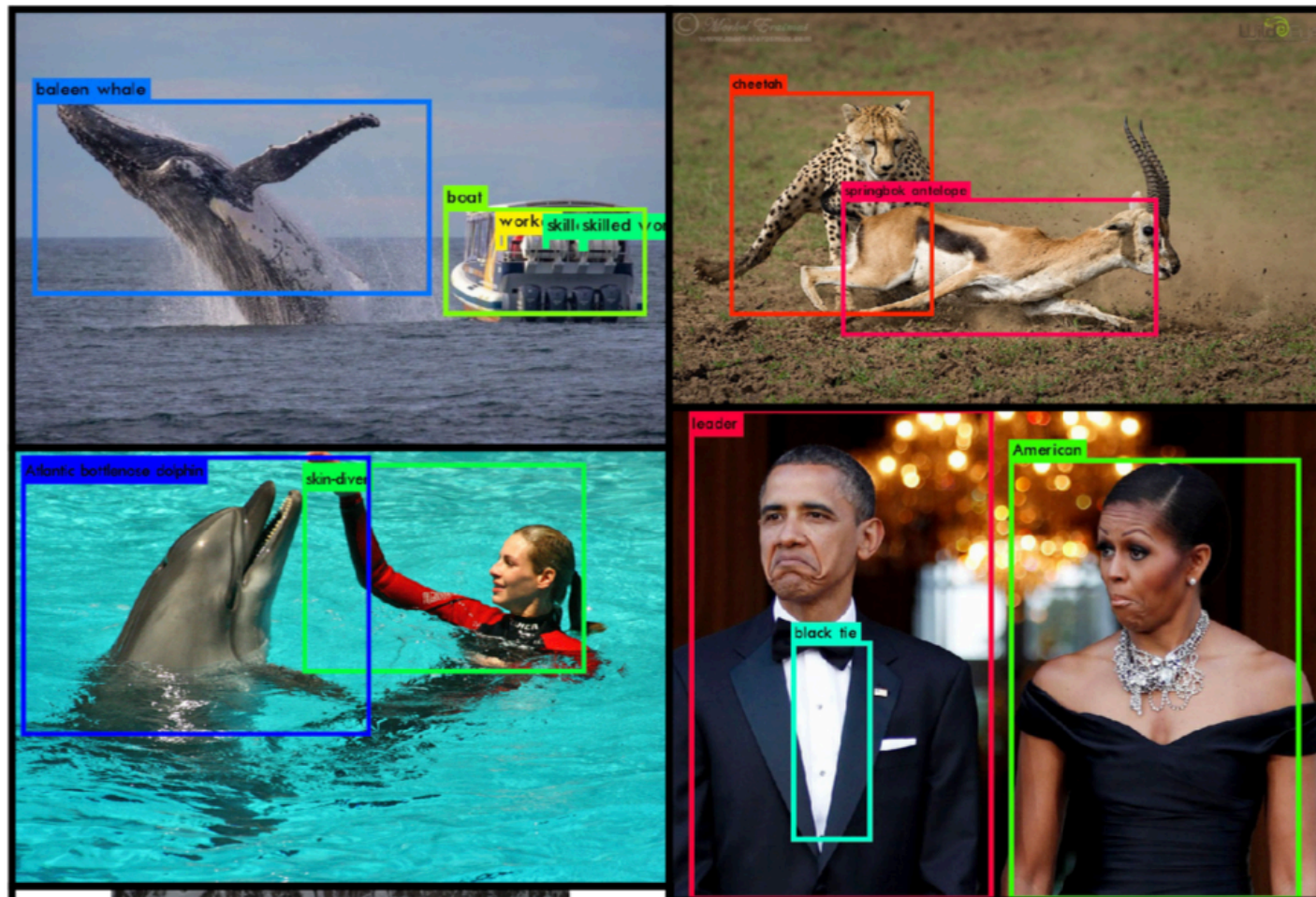
Object detection with *Detectron*:



Source: R. Girshick, I. Radosavovic, G. Gkioxari, P. Dollár, und K. He, Detectron. 2018.

System - Frameworks - Yolo

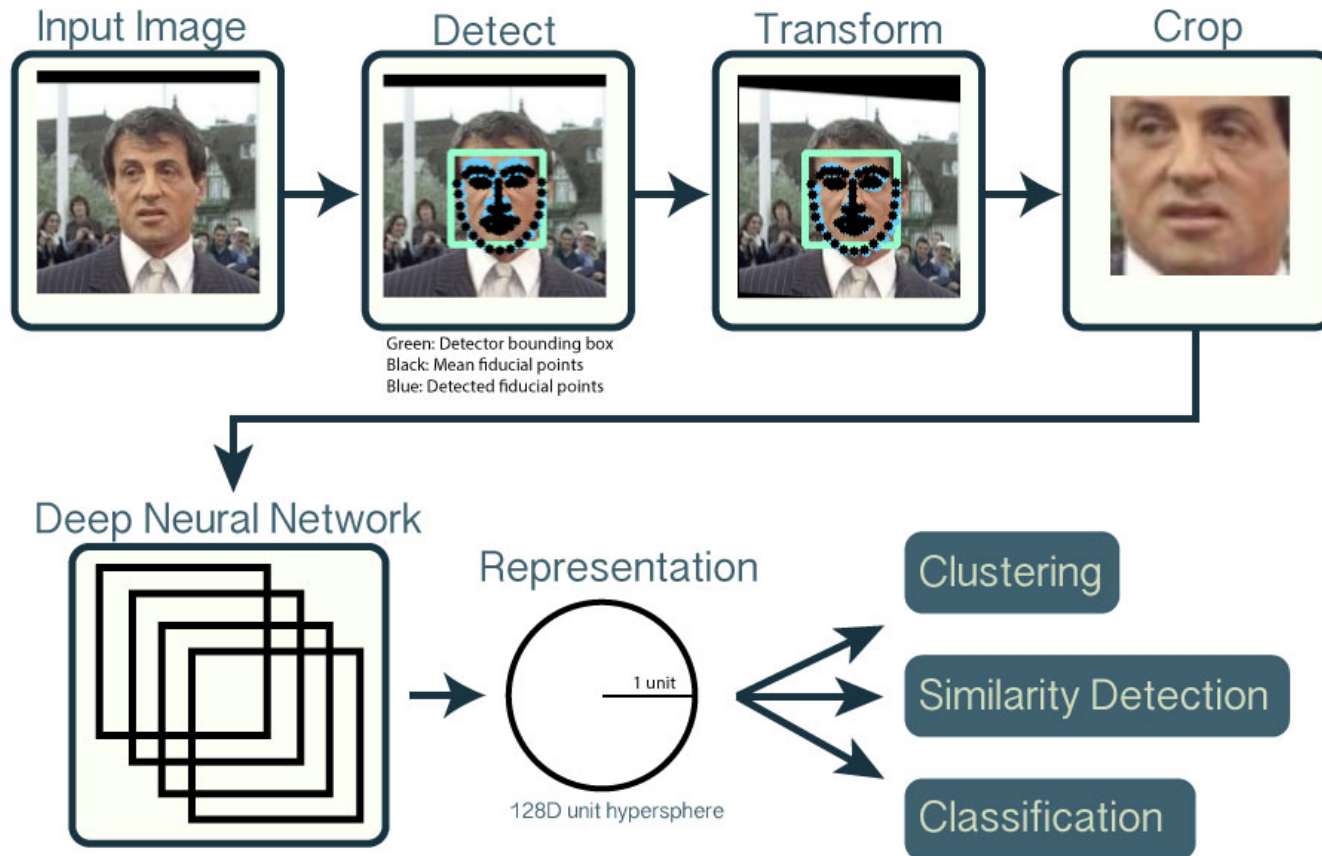
Real-time object detection with *Yolo9000*.



Source: Joseph, R. and Ali, F., 2016. Yolo9000: Better, faster, stronger. arXiv preprint arXiv:1612.08242

System - Frameworks - OpenFace

Face recognition with *OpenFace*:



Source: Amos, Brandon and Bartosz Ludwiczuk and Satyanarayanan, Mahadev, OpenFace: A general-purpose face recognition, 2016

System - Frameworks - Face Recognition

Detect and recognize faces with *Face Recognition*:



Input

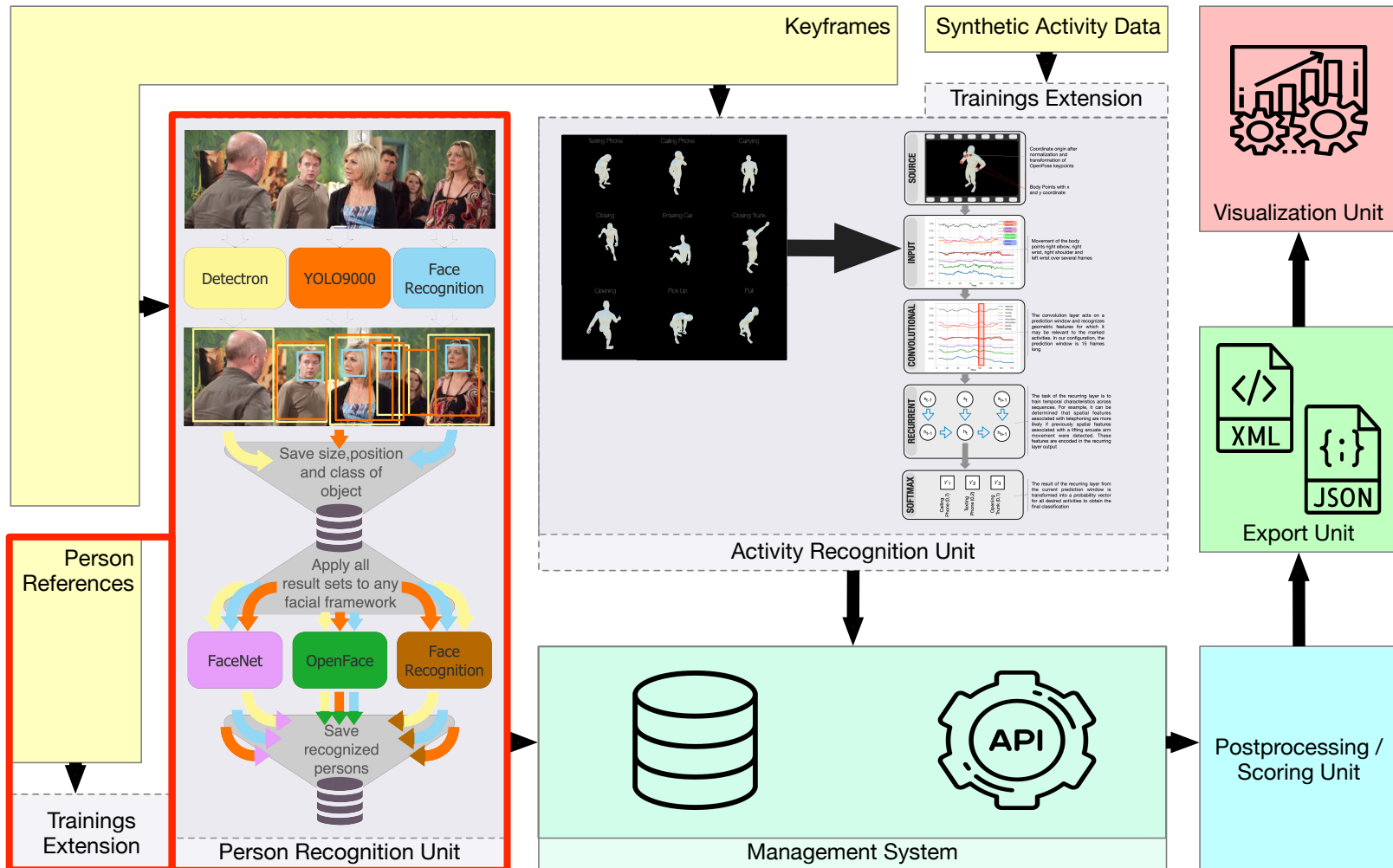


Output

Simplify the development of custom machine learning models with Turi Create:

ML Task	Description
Recommender	Personalize choices for users
Image Classification	Label images
Drawing Classification	Recognize Pencil/Touch Drawings and Gestures
Sound Classification	Classify sounds
Object Detection	Recognize objects within images
One Shot Object Detection	Recognize 2D objects within images using a single example
Style Transfer	Stylize images
Activity Classification	Detect an activity using sensors
Image Similarity	Find similar images
Classifiers	Predict a label
Regression	Predict numeric values
Clustering	Group similar datapoints together
Text Classifier	Analyze sentiment of messages

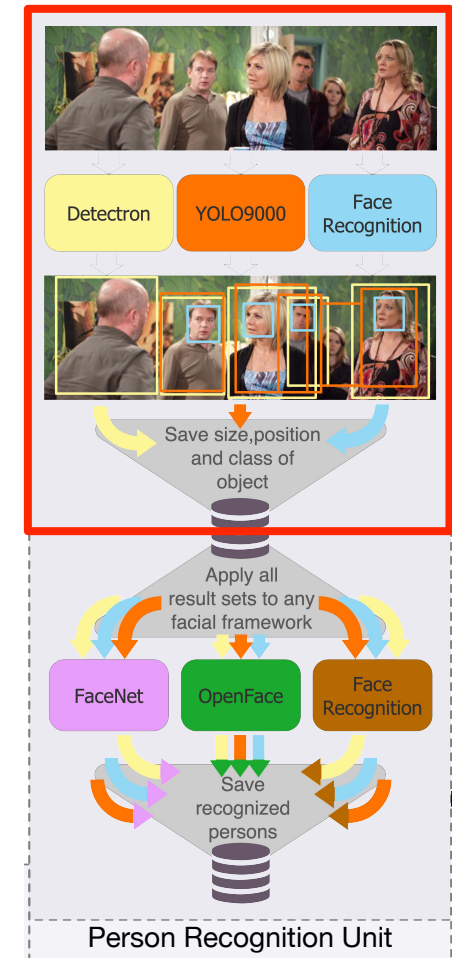
Workflow - Person Recognition Unit



Workflow - Person Recognition Unit

Person recognition in three steps

- Localization of the absolute position
 - Detectron 2.515.332 person objects
 - Face Recognition 1.384.747 faces
 - YOLO9000 1.013.007 person objects
 - Each results stored separate in database table

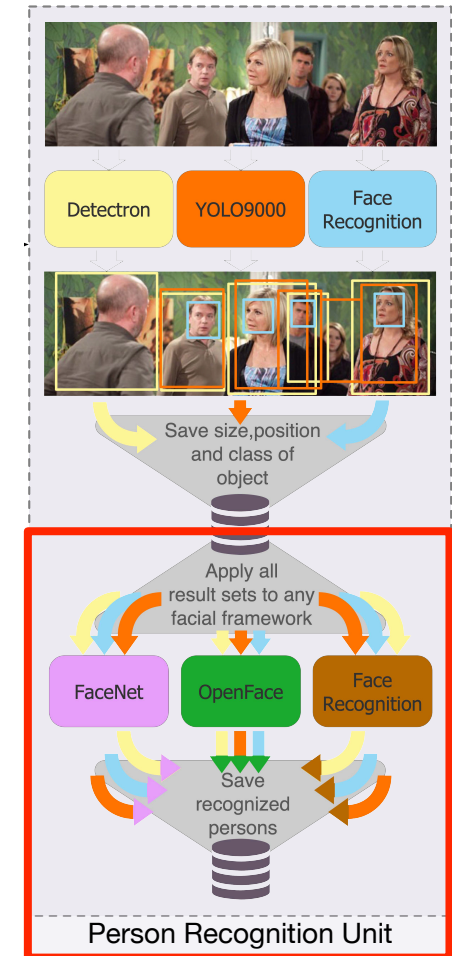


Workflow - Person Recognition Unit

Person recognition in three steps

- Localization of the absolute position
 - Detectron 2.515.332 person objects
 - Face Recognition 1.384.747 faces
 - YOLO9000 1.013.007 person objects
 - Each results stored separate in database table

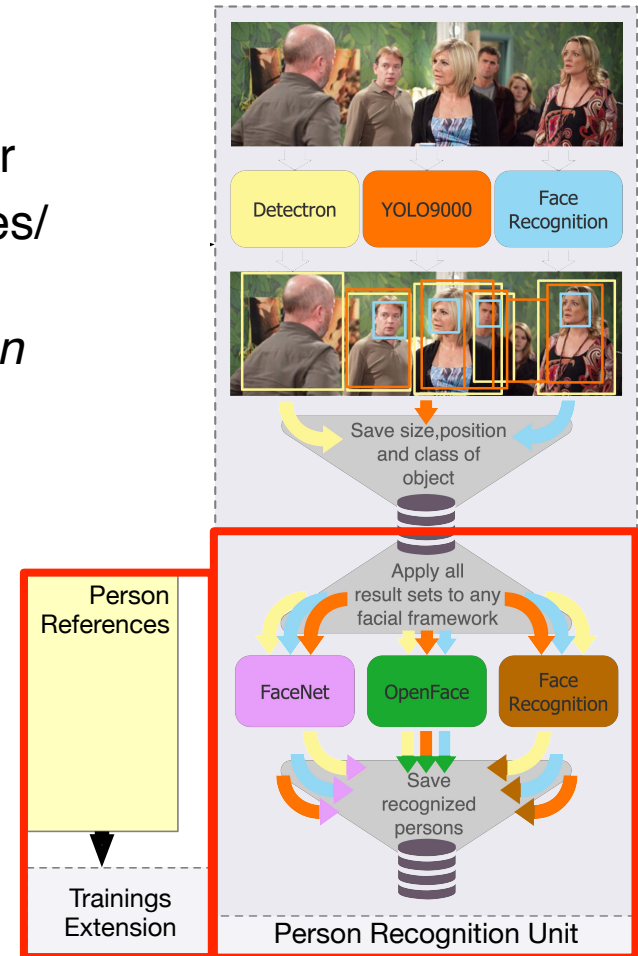
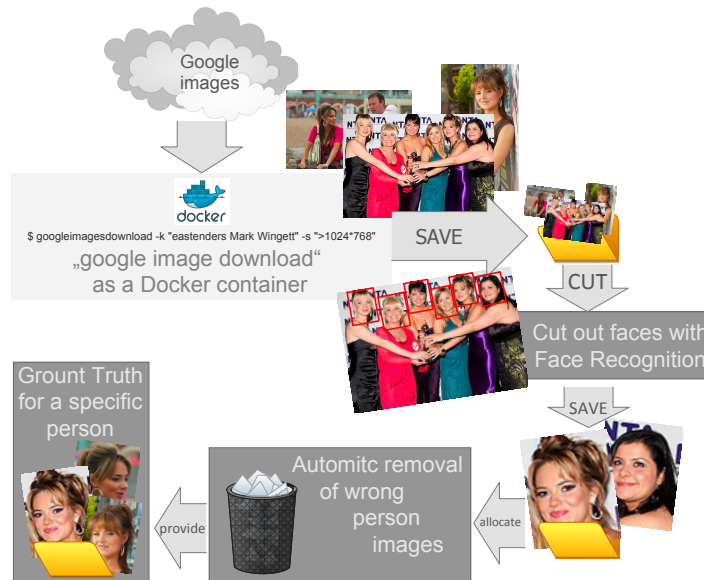
- Person recognition
 - Classification of the persons
 - Usage of models trained on dataset collected with google-image-downloader



Workflow - Person Recognition Unit

Person recognition in three steps

- Create ground truth data
 - Usage of google-image-downloader
 - Collection an average of 150 images/ person
 - Cut out faces with *Face Recognition*
 - Train person models with three frameworks



Workflow - Person Recognition Unit

Person recognition in three steps

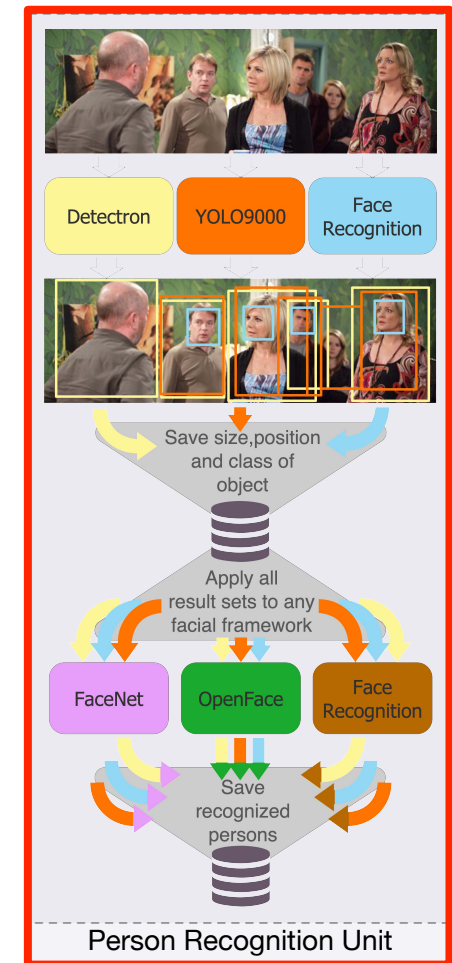
- Merging the results of all frameworks to a person score

$$predPerson = \sum_{h=0}^r (x_h \in K), \quad (1)$$

$$K = \{x | 0 \leq x \leq 100\}, \quad (2)$$

$$predPerson \in L, \quad (3)$$

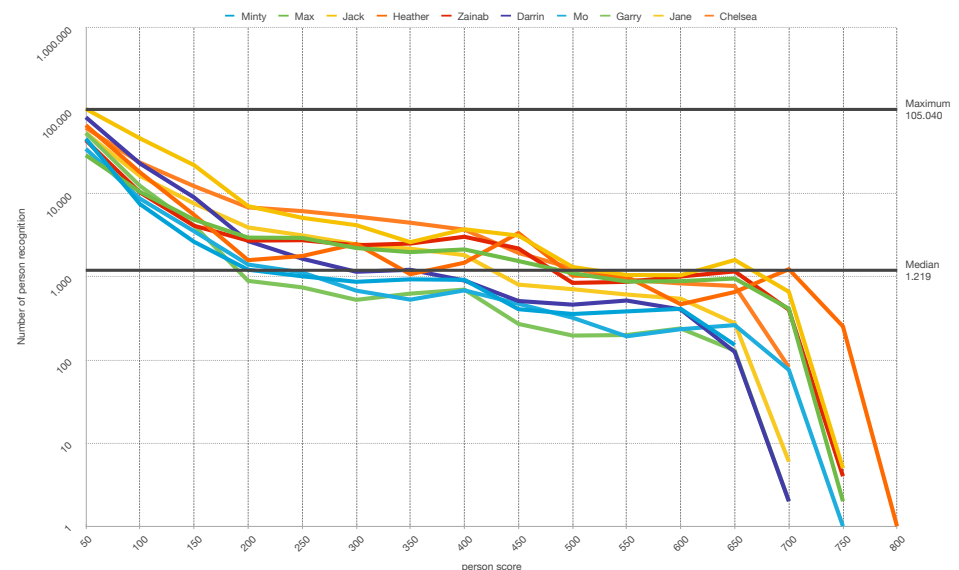
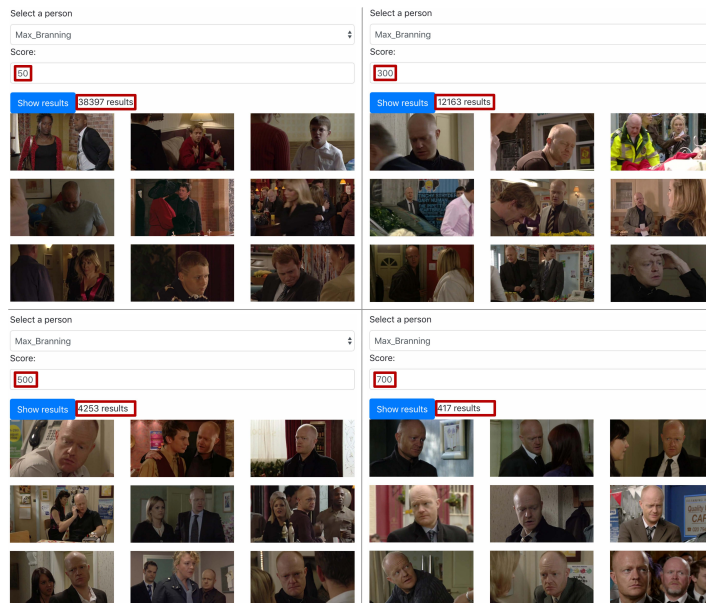
$$L = \{y | 0 \leq y \leq 900\}. \quad (4)$$



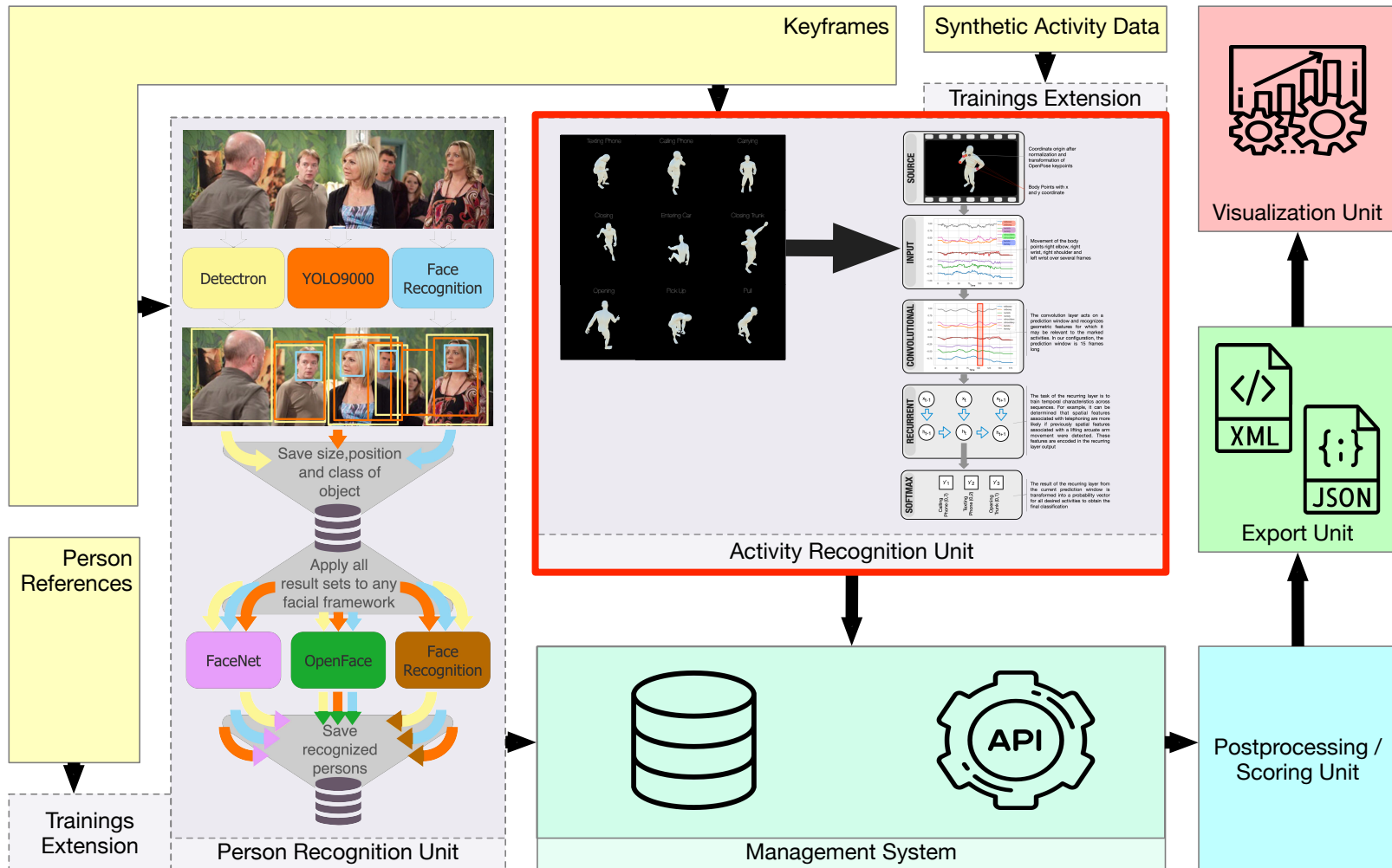
Workflow - Person Recognition Unit

Scoring Results

- We developed 2018 a web-service for interactive evaluation
- The findings from the visual processing not include in the automatic evaluation
- Number of fail detections decrease with increasing score



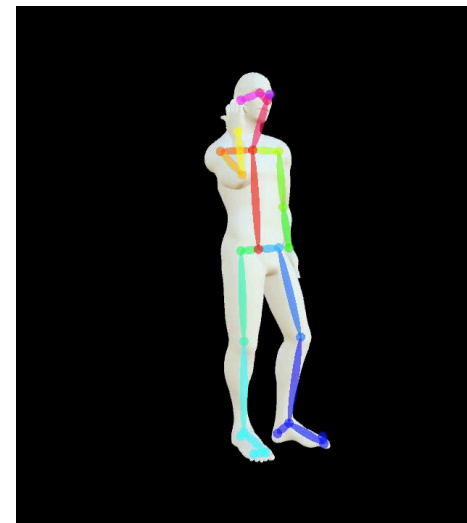
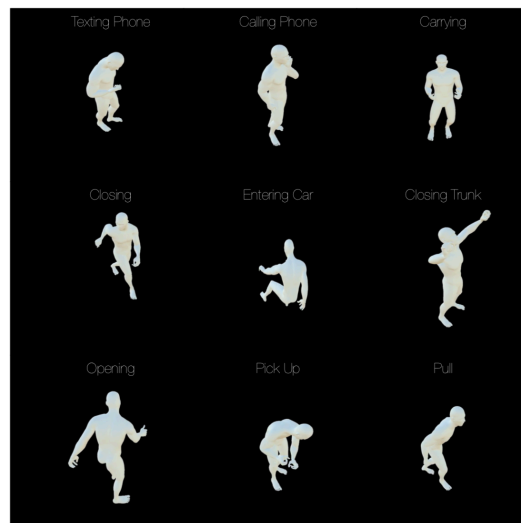
Workflow - Activity Recognition Unit



Workflow - Activity Recognition Unit

Different approaches for activity recognition

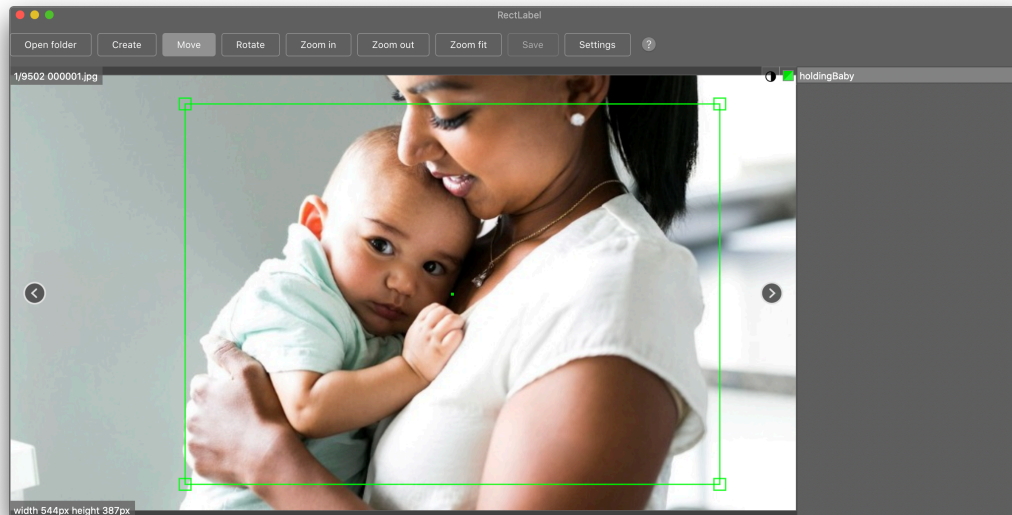
- Self-developed activity classifier
 - Create our own suitable synthetic ground truth dataset
 - Extracting body-key-points based on the results by *OpenPose*
 - Able to predict activities in realtime



Workflow - Activity Recognition Unit

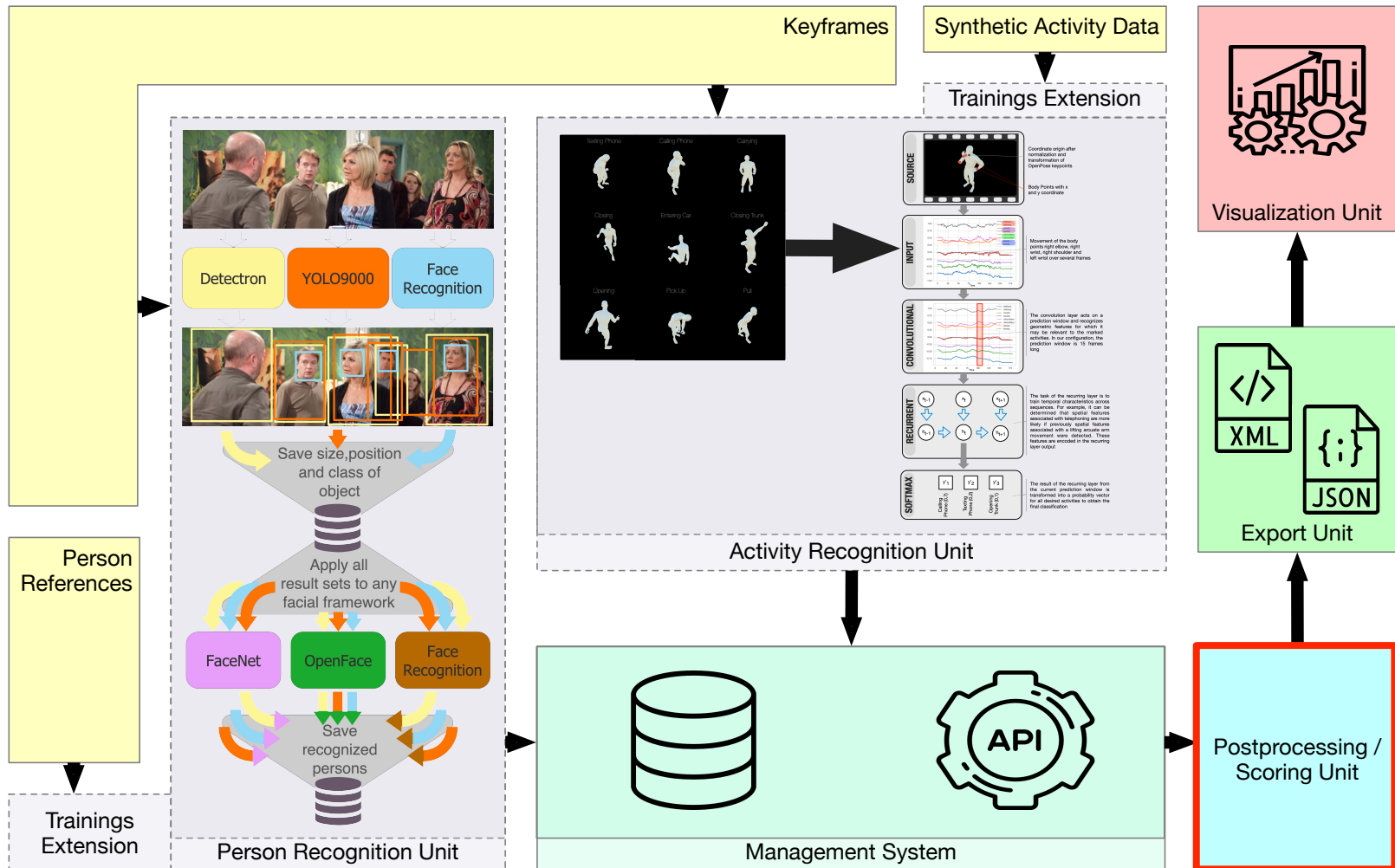
Different approaches for activity recognition

- Object Activity classifier (Detectron, YOLO9000, Custom-Model)
 - The classifier was trained with 9504 images showing drinking, eating, holdingBaby, holding- Glass, holdingPhone, hugging, laughing and kissing
 - Images are downloaded with *GoogleDownloader* and manually labelled using “*RectLabel*” software
 - The object recognition classify and localize the activities



Source: Screenshot macOS app „RectLabel“

Workflow - Scoring Unit



Workflow - Scoring Unit

Fusion of the determined scoring values

- The fusion took place exclusively in the database
- Linked the various framework results with suitable SQL statements
- An overall score is be calculated with the person and activity recognition results

Conclusion

Our INS approach with pretrained models and web based interactive evaluation by using a management system

- The system compute twelve times more frames as last year in the same time
- So we significantly improved performance
- The using of the distributed approach lets us calculate parallel on several machines
- A functional optimization of the newly developed parts follows next year
- So an increase in result quality after refinement is expected
- This was a practical evidence, that our optimized application is able to handle the provided data in an acceptable computational time
- With the focus on high adaptability the system could be extended every time by frameworks of other application domains